Immunity, Vol. 7, 13–24, July, 1997, Copyright ©1997 by Cell Press

Complementation of V(D)J Recombination Deficiency in RAG-1$^{-/-}$ B Cells Reveals a Requirement for Novel Elements in the N-Terminus of RAG-1

Christopher A. J. Roman, Sara R. Cherry, and David Baltimore
Department of Biology
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Summary

RAG-1 is an essential component of the site-specific V(D)J recombinase. A new assay system has revealed a significant contribution of the catalytically dispensible N-terminal region of RAG-1 to recombination activity. The foundation for this system is an Abelson virus–transformed cell line derived from RAG-1$^{-/-}$ mice that is dependent on the introduction of exogenous RAG-1 for rearrangement of either plasmid substrates or the endogenous immunoglobulin loci. Use of this line demonstrates that conserved and novel cysteine-containing elements in the N-terminal region are required for full RAG-1 activity when recombination activity is in a RAG-1 dose-responsive range. Our data suggest that the RAG-1 N-terminus enhances the formation of an active recombination complex that facilitates the rearrangement process.

Introduction

The recombination activating gene-1 (RAG-1) and RAG−2 proteins are essential for mediating the somatic assembly of antigen receptor genes in a process known as V(D)J recombination (Gellert, 1996; Oettinger, 1996). Sequences flanking V, D, and J gene segments that act as target sites for V(D)J recombination are composed of a conserved heptamer and nonamer separated by 12 or 23 nucleotides and are known as recombination signal sequences (RSSs). V(D)J recombination first involves the generation of double-strand DNA breaks at RSSs by the RAG proteins, followed by the reciprocal exchange, modification, and ligation of the broken DNA ends, which require proteins involved in double-strand-break repair (Hendrickson et al., 1991; Jackson and Jeggo, 1995; Weaver, 1995). The products are signal joints (from the normally perfect joining of the blunt heptamer ends) and coding joints (from the joining of the gene segment ends via hairpin intermediates).

Catalytic "core" domains of RAG-1 (amino acids [aa] 384–1009) and RAG-2 (aa 1–387) have been defined by genetic and biochemical means. These core segments of the proteins mediate many activities associated with V(D)J recombination. For example, they are sufficient to recombine episomal and integrated-retroviral substrates when coexpressed in fibroblasts (Oettinger et al., 1990; Sadofsky et al., 1993; Silver et al., 1993; Kirch et al., 1996). In vitro, the core proteins together cleave DNA at heptamer/coding end borders in an RSS-dependent (McBlane et al, 1995; Cortes et al., 1996) and 12/23-rule–dependent manner (Eastman et al., 1996; van Gent et al., 1996b), with artificial substrates. It has been shown that the core domain of RAG-1 alone can recognize the RSS (Spanopoulou et al., 1996) and binds to the nonamer element via a characteristic homeodomain (Difilippantonio et al., 1996; Spanopoulou et al., 1996).

The present study arose from two interests: exploration of the role the catalytically dispensible portions of RAG-1 might play in V(D)J recombination and of whether measurement of V(D)J recombination in lymphocytes using endogenous genes as substrates might reveal important but previously unexamined aspects of the process. The major catalytically dispensible portion of RAG-1 is its N-terminal region, which has features suggesting that it should have functional importance. It constitutes one third of the protein (aa 1–383) and has been conserved and maintained throughout vertebrate evolution (Figure 1; Sadofsky et al., 1993; Willett et al., 1997). A well-conserved motif within this region is a zinc-binding RING finger (aa 290-338; Saurin et al., 1996) that can form homodimers (Rodgers et al., 1996). Clusters of basic amino acids within the N-terminal region interact with the nuclear importin hSRP and are required for the nucleolar localization of RAG-1 (Cortes et al., 1994; Spanopoulou et al., 1995). Collectively these observations suggest that this region may have a role in regulating RAG-1 activity.

Most assays of V(D)J recombination in cells have employed artificial substrates and nonlymphoid cells. Our second goal was to establish a more physiological assay, in which RAG-1 activity could be measured both within a cellular milieu in which V(D)J recombination normally occurs and by its ability to rearrange antigen–receptor gene segments within their normal chromosomal context. Here we describe a system in which we complement the recombination deficiency of a RAG-1$^{-/-}$ Abelson virus–transformed line by the transient expression of exogenous RAG-1. An important feature of this system is that it facilitates analysis of many different alleles of RAG-1 for their ability to rearrange both plasmid and endogenous immunoglobulin targets without an extended selection process and the biases that selection introduces.

In the Abelson line, we found that truncations within the RAG-1 N-terminal region had a pronounced detrimental effect on recombinational activity measured with plasmid substrates, distinctly greater than that previously observed in fibroblasts (Sadofsky et al., 1993; Silver et al., 1993; Kirch et al., 1996). The behavior of the truncations allowed us to delineate novel and conserved cysteine elements the integrity of which is required for full RAG-1 activity in this system. We also found that the homologous domain from the evolutionarily distant zebrafish RAG-1 protein can provide the activity. A critical parameter that allowed us to reveal the activity of the N-terminal region is that the Abelson line provided a milieu in which recombination activity is in a RAG-1 dose-responsive range; previous systems have not had that property. The system also allowed us to show that in RAG-1 excess, both the core domain of murine RAG-1 and full-length zebrafish RAG-1 can mediate endogenous immunoglobulin V(D)J recombination.

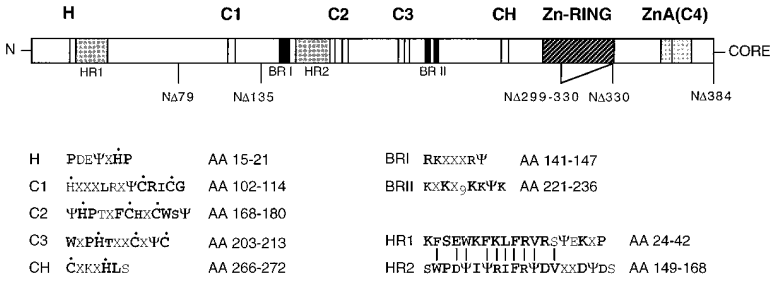

Figure 1. Schematic Representation of the Conserved Elements within the N-Terminal 384 Amino Acids of Murine RAG-1

H, C1, C2, C3, CH, Zn-RING and ZnA are conserved cysteine/histidine elements in the N-terminus. The Zn-RING motif and ZnA have been described previously (Rodgers et al., 1996) and their labels are positioned over shaded regions, which span the conserved residues. H, C1, C2, C3, and CH are positioned over lines indicating the conserved residues. BRI and BRII are BRs that can act as nuclear localization signals (Spanopoulou et al., 1995). HR1 and HR2 are conserved blocks of amino acids that share sequence homology with each other (shown by lines between the sequences). The amino acid positions that demark the truncations of RAG-1 discussed in the text are listed below the diagram. The sequences of the conserved elements and their amino acid (AA) positions are indicated. The extent of amino acid conservation here is based on the comparison of RAG-1 proteins described by Willett et al. (1997) and also includes comparison to the shark (Bernstein et al., 1996) and rabbit (Fuschiotti et al., 1993) RAG-1 sequences. H and HR1 are not obviously present in shark RAG-1. With that exception, amino acids present in all species are indicated by large boldface letters. Ψ indicates that an aliphatic or aromatic residue is present at that position in all species; small boldface letters indicate absolute conservation of related amino acids (i.e., I and L, D and E, R and K, and S and T); small lightface letters indicate a consensus amino acid but one that is not conserved with homologous amino acids in all species; X indicates no obvious amino acid preference.

Results

Truncated Forms of RAG-1 Do Not Efficiently Complement the Recombination Defect in RAG-1$^{-/-}$ Abelson Lines with Exogenous Plasmid Substrates

v-*abl*–transformed (Abelson) B cell lines have historically provided an important model system for the study of V(D)J recombination (Rosenberg et al., 1975). Abelson lines from wild-type mice express both RAG-1 and RAG−2 proteins and normally undergo immunoglobulin recombination in culture (Alt et al., 1981). To establish a system in which the activity of wild-type and mutated alleles of RAG-1 could be tested in a physiological context, we examined whether Abelson lines from the bone marrow of adult RAG-1$^{-/-}$ mice had such properties. As expected, all immunoglobulin loci were in germline configuration in these lines, and they were devoid of V(D)J recombination activity (see below). A more detailed phenotypic analysis showed that they were identical to their wild-type counterparts in terms of cell surface marker expression and the expression of a panel of B cell specific genes, including RAG-2 and germline immunoglobulin heavy-chain transcripts (C. A. J. R., unpublished data).

We chose one line (1-2) to test whether the lack of V(D)J recombination activity could be restored by the addition of exogenous wild-type RAG-1 (herein referred to as R1WT) by diethylaminoethyl (DEAE)–dextran–mediated transient transfection of DNA encoding the protein. Two days after transfection of 1-2 cells with the RAG-1 expression vector CDM8-R1WT and either recombination substrate pJH289 or pJH288 (Lieber et al., 1988), a total cellular lysate was analyzed using a polymerase chain reaction (PCR) assay designed to amplify a DNA fragment dependent on the formation of a signal joint by V(D)J recombination (Roman and Baltimore, 1996). Transient expression of R1WT in the Abelson line activated the recombination of both plasmid substrates (Figure 2A, compare lanes 1 and 5 and lanes 10 and 11). Therefore, supplementing the 1-2 cells with R1WT alone was sufficient to complement the plasmid V(D)J recombination deficiency.

Previous genetic and biochemical studies have defined a minimal catalytic "core" of murine RAG-1, aa 384–1009, that is essential for V(D)J recombination. In fibroblasts, deletion of the N-terminus of RAG-1 resulted in only a 2- to 3-fold loss in activity as measured with artificial substrates (Sadofsky et al., 1993; Silver et al., 1993; Kirch et al., 1996; see also below). However, when we tested truncations of RAG-1 (Figure 1; Silver et al., 1993) for their efficacy in activating recombination of plasmid substrates by transient transfection into the RAG-1$^{-/-}$ Abelson line, their activity was generally 10- to 20-fold diminished compared to R1WT (Figure 2A, compare lanes 7, 9, 13, and 15 with lanes 1–5, and Figure 2B).

Novel and Conserved Cysteine Elements in RAG-1 Are Critical for Its Activity in the Abelson Line

In the comparison of deletion mutants, there was a large difference in activity when the deletion was extended from NΔ13–79 to NΔ13–135, the larger deletion behaving like the more severely truncated proteins NΔ13–330 and NΔ1–384 (Figure 2A, lanes 6, 7, 9, 12, 13, and 15, and Figure 2B). The extended deletion spans a region containing the sequence CR(I/L)CG (C110–G114), which is conserved in all RAG-1 proteins for which sequence data has been obtained. The CRICG element resembles the C_2 portion of C_2H_2 zinc fingers and nucleocapsid proteins; indeed, a BLAST search (Altschul et al., 1990) showed that these amino acids share sequence homology to the C_2 of several zinc fingers, a subset of which are shown in Figure 3A. Furthermore, seven of eight RAG-1 proteins contain a histidine residue (H102) 8 amino acids N-terminal of the first cysteine. Interestingly, a histidine residue not part of an N-terminal zinc finger often appears with the same spacing before the first cysteine residue of C_2H_2 zinc fingers within a number of DNA-binding proteins (Figure 3A, asterisks).

A comparison of RAG-1 proteins from different organisms showed that several other cysteine and histidine clusters were absolutely conserved (Figure 1). In addition to those in the Zn-RING, these are H170C175C178 (C2), H206C210C213 (C3), C266H270 (CH), and

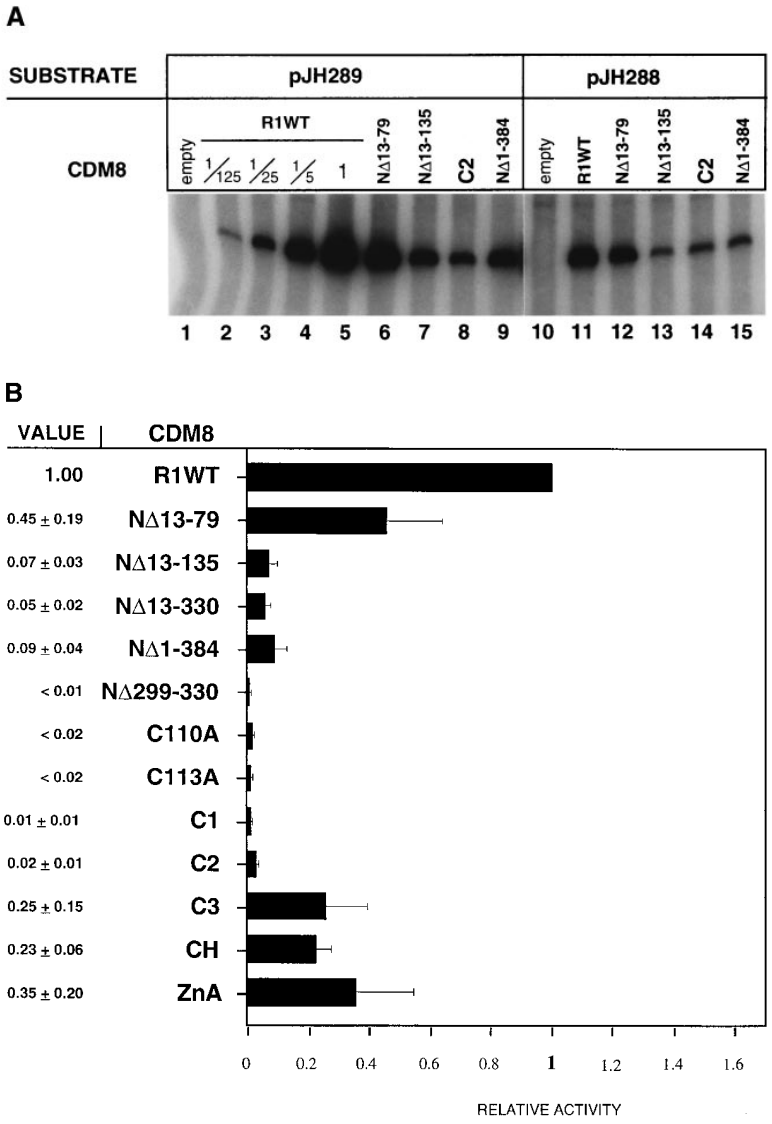

Figure 2. Activation of Episomal V(D)J Recombination in the RAG-1$^{-/-}$ Abelson line 1-2 by Exogenous RAG-1 Alleles

(A) Autoradiograph of plasmid recombination PCR fragments derived from appropriately rearranged pJH288 and pJH289 substrates. The fragments were amplified from lysates of 1-2 cells transiently transfected with the recombination substrate and CDM8-driven full-length or truncated RAG-1 cDNAs. The RAG-1 allele transfected is indicated above each lane. For the dilution samples, "R1WT pJH289" lysate was diluted into "empty" lysate and amplified in parallel; the number above the lane indicates the fraction of R1WT pJH289 present in the sample. The values for relative activity was determined from the dilution series.

(B) Bar graph indicating the relative activities of truncated and point-mutated RAG-1 alleles in the Abelson line 1-2. The activity of R1WT is defined as 1.00 and is directly proportional to the intensity of the recombination PCR product as determined by phosphorimaging of the radioactive bands (see Experimental Procedures).

C355C360H372 (ZnA). The ZnA (Rodgers et al., 1996) and the Zn-RING together mediate homodimerization. In addition to the conserved CxxC element, C1, C2, and C3 all had a conserved N-terminal histidine residue (Figure 3B). Furthermore, conserved acidic residues were present upstream of the histidines; interestingly, in addition to histidine and cysteine, acidic residues have also been shown to participate in coordinating zinc (Cunningham et al., 1991). Therefore, elements C1, C2, and C3 share the same panel of amino acids implicated in coordinating zinc ions in other contexts. In contrast, in CH we found conserved acidic residues C-terminal to the conserved histidine and cysteine, but the spacing of the conserved residues was similar to those in C1, C2, and C3. Although each cysteine element contains idiosyncratic conserved amino acids, they contain hallmark features that support the contention that they have a metal coordinating role. We named this conserved set of amino acids a D/EHC$_2$ element.

To determine whether the conserved two cysteines in element C1 were critical to RAG-1 activity, we constructed mutants C110A and C113A and combined them in mutant C1. Relative to the wild type, these point mutations gave an even lower activity than the deletion in NΔ13–135 (Figure 2B and Figure 3C, compare lanes 3–5 with lanes 2 and 15–17), showing their importance for RAG-1 activity in this system. We then mutated the conserved cysteine and/or histidine residues to alanines within the C2, C3, CH, and ZnA elements and tested the mutant proteins in 1-2 cells. Of these mutants, only C2 had as extreme an effect as the C1 mutant, although all had lower activity than R1WT (Figure 2A, lanes 8 and 14; Figure 2B; and Figure 3C, lanes 6–9). The mutagenic data show that the C1 and C2 elements are crucial to RAG-1 activity and the C3, CH, and ZnA elements also play a role. These are all likely to be metal coordination elements and suggest a complex organization for the N-terminal domain.

Other point mutations in the N-terminal region of RAG-1 did not have as strong effects on recombination activity. These mutations included ones in the basic regions (BRs), which are important for mediating nuclear and nucleolar localization (Spanopoulou et al., 1995). We observed that although some mutations in the basic

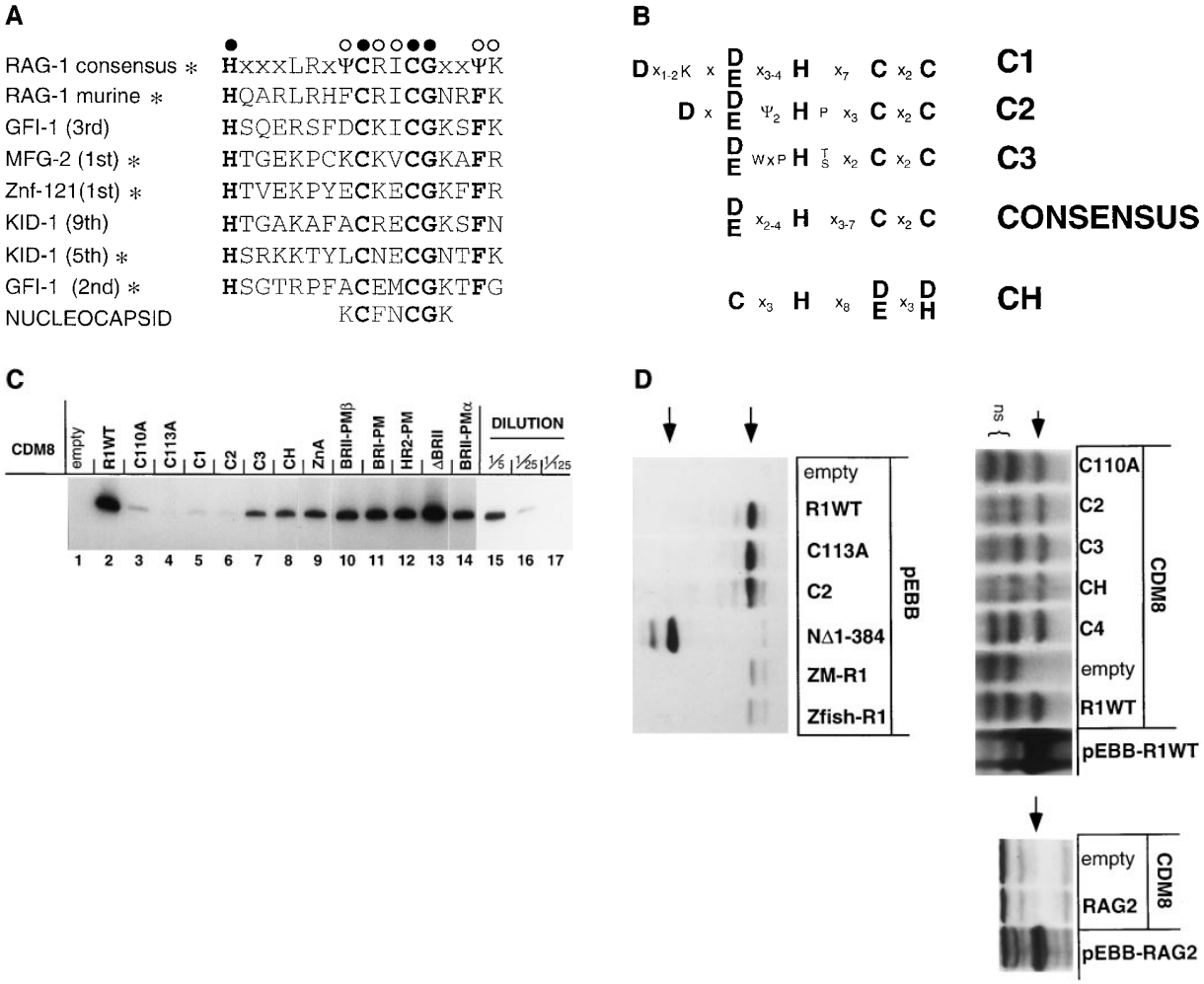

Figure 3. Conserved Cysteine/Histidine Elements in the N-Terminus of RAG-1 Are Required for RAG-1 Activity in the Abelson Line

(A) Sequence alignment of the C1 element with the CxxC portion of the C_2H_2 zinc-finger proteins GFI-1 (Gilks et al., 1993), MFG-2 (Passananti et al., 1989), Znf-121 (Chen et al., 1993), and KID-1 (Witzgall et al., 1993), and the HIV nucleocapsid protein (Myers et al., 1993). An asterisk indicates that the sequence is taken from a zinc finger that lacks another finger immediately N-terminal to it; therefore in those cases the first histidine residue is not part of a preceding zinc finger. In the remainder, the histidine residue is the last histidine residue in the preceding finger. The ordinal number in parentheses indicates the placement of the finger of a series in the protein. In the RAG-1 consensus sequence, putative coordinating amino acids and the absolutely conserved glycine residue are shown in boldface; other consensus amino acids are shown in lightface. In this case, the amino acid consensus does not represent absolute conservaton as in Figure 1. A filled circle over the amino acid indicates that the residue in the RAG-1 consensus is present in all sequences listed; an open circle indicates that is present in a subset.

(B) Sequence comparison of the C1, C2, C3 and CH elements. For C1, C2, and C3, the CxxC elements were used to align the other amino acids. Large capital boldface letters indicate conserved and putative zinc-coordinating amino acids (D/E and H). Conserved amino acids in each individual element are indicated by small lightface letters, and the amino acid notation is as in Figure 1. Subscript numerals below each X indicate the number of amino acids represented by the X. The following are the specific exceptions to the consensus: in C1, trout contains a tyrosine at the H position, and human a glycine at the D/E position, and shark lacks an obvious D, K, and D/E consensus amino acid; in C2, shark contains a cysteine at the H position; in C3, trout has a glycine at the D/E position; and in CH, shark has a glutamine at the D/H position.

(C) Autoradiograph of plasmid recombination PCR products from 1-2 cells transfected with CDM8-driven wild-type and point-mutated RAG-1 cDNAs. C110A and C113A are point mutants of the two conserved cysteine residues in C1; HR2 contains point mutations within that conserved element; BRI and II indicate mutations within Basic Regions described in Experimental Procedures. PM, a point mutation; Δ, a deletion.

(D) Western blot analysis of RAG proteins used in this study. CDM8- or pEBB-RAG-1 alleles were transiently expressed in the cell line 293. Total cell lysates were resolved by SDS–polyacrylamide gel electrophoresis, and proteins transferred to nitrocellulose. (Top left) A mouse anti-murine RAG-1 monoclonal antibody was used to detect RAG-1 proteins expressed from pEBB. (Top right) A rabbit anti-RAG-1 polyclonal was used to detect RAG-1 proteins expressed from CDM8. (Bottom) A mouse anti-murine RAG-2 monoclonal was used to detect RAG-2 protein. Arrows indicate the positions of the specific bands; ns indicates nonspecific bands in the cell lysate, which reflect total protein and loading efficiency.

regions had 2- to 3-fold less activity than R1WT (Figure 3C, lanes 10, 11, 13, and 14), there was no consistent correlation between the ability of RAG-1 proteins mutated in these regions to activate recombination and their nuclear localization (Figure 3C and data not shown). It is possible that the nuclear localization properties of the BRs in the N-terminus are redundant to those in the C-terminus. In addition, a point mutation within the internal homology region 2 (HR2) (Figure 1) also had little effect (Figure 3C, lane 12), and a truncation of the C-terminal 17 amino acids slightly activated recombination (data not shown). These results further emphasize that the severe effect of mutations in C1 and C2 was due to the loss of a property specific to the cysteine-containing elements.

We also tested the internal deletion mutant NΔ290–330, which contains a deletion of most of the Zn-RING. This mutant had even lower activity than the complete N-terminal deletion mutant NΔ1–384 (Figure 2B). However, this RAG-1 allele contains the C1 and C2 elements; its low activity therefore suggested that C1 and C2 may not be functionally autonomous elements. Conversely, the mutant NΔ13–135 contains the Zn-RING and ZnA yet also lacks full RAG-1 activity. Therefore, all of the cysteine-containing elements are required to maintain the full functional integrity of the N-terminal region as demonstrated in 1-2 cells, although a more specific structural relationship between them cannot be specified.

A Western blot analysis of the RAG-1 mutant proteins expressed in 293 cells showed that the steady-state amounts of mutant proteins was comparable with R1WT (Figure 3D; Silver et al., 1993). Furthermore, immunofluorescent localization of the protein in 293 cells transfected with the C2 mutant showed it to be in the nucleoli, indistinguishable from R1WT localization (data not shown). It is formally possible that cell line–specific differences in protein stability or localization may exist, but we have not been able to detect RAG-1 proteins from 1-2 cells transiently transfected by DEAE-dextran. However, it has been reported that epitope-tagged, truncated RAG-1 proteins have been detected in a wild-type Abelson line after electroporation and that they were expressed at wild-type levels in the nuclear fraction (Sadofsky et al., 1993).

The Core Domain of RAG-1 Is Sufficient for Mediating Immunoglobulin Chromosomal Recombination In Vivo When Expressed at High Levels

In the 1-2 line, the immunoglobulin heavy-chain locus is transcriptionally active and should be poised for recombination but for the lack of RAG-1. We thus tested whether transiently expressed RAG-1 could activate the recombination of the immunoglobulin heavy-chain locus in these cells. For the following experiments, we used the vector pEBB (Mayer et al., 1995) containing the E1α promoter to drive maximal expression of the RAG-1 alleles. RAG-1 protein was expressed at approximately 10- to 20- fold higher amounts with the pEBB vector than with CDM8 (Figure 3D, top left, last two lanes). To monitor the activation of endogenous V(D)J recombination, we prepared genomic DNA from 1-2 cells transiently transfected with pEBB-R1WT and used a PCR assay to detect DJ_H specific rearrangements (Schlissel et al., 1991; Zhu et al., 1996). Wild-type RAG-1 promoted the formation of DNA fragments that corresponded to those found in vivo (Figure 4A, lane 2). Therefore, RAG-1 was the only missing component required to activate immunoglobulin recombination in 1-2 cells. Most significantly, we could also detect the production of all four DJ_H rearrangement products using the deletion mutant NΔ1–384 (Figure 4A, lane 7). We can thus conclude that the N-terminal 383 amino acids do not contribute a unique mechanistic function necessary for mediating chromosome recombination in vivo.

The formation of DJ_H products was also promoted by other RAG-1 deletion and point mutants, although at much lower frequencies (Figure 4A, lanes 4–6). Usually, only a subset of fragments was amplified during any one PCR assay. This was most likely due to the low frequency of rearranged alleles in the cell population, leading to detection of only some bands in each PCR assay (Chang et al., 1992). We could collectively visualize all four rearrangement products from separate PCR

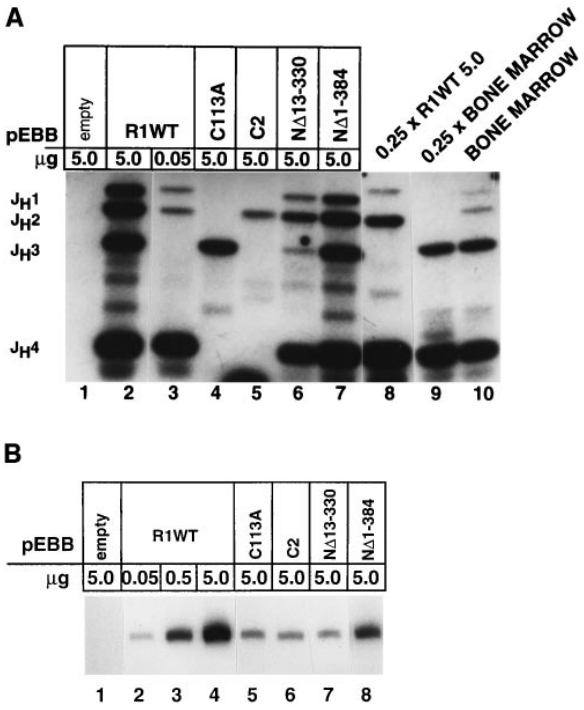

Figure 4. Transient Transfection of pEBB-RAG-1 Alleles into the RAG-$1^{-/-}$ Abelson Line 1-2 Activates Endogenous D_H-to-J_H Recombination

(A) Autoradiograph of DJ_H PCR fragments amplified from genomic DNA of 1-2 cells transiently transfected with different RAG-1 alleles. The identity and the amount of transfected vector are indicated above each lane. "5.0 μg R1WT" DNA was diluted 4-fold into "empty" DNA to generate the sample "0.25 × R1WT 5.0" for DJ_H PCR amplification. A sample of DNA from wild-type murine bone marrow and a 4-fold (0.25×) dilution of it are included as positive controls. The identity of each PCR product is indicated adjacent to the autoradiograph.

(B) Autoradiograph of pJH289 recombination PCR products amplified from 1-2 cells transfected with the indicated pEBB-RAG-1 allele. The amount transfected is also indicated. A bar graph comparing the relative activities of some of the RAG-1 alleles expressed from pEBB in 1-2 cells is shown in Figure 5A.

amplifications of these samples run in parallel (S. R. C., unpublished data). We also observed that transfection of less R1WT, for instance using less pEBB-R1WT or with the CDM8 vector, also decreased the amount of DJ_H amplification products and the frequency at which specific amplified products were detected (Figs. 4A lane 3 and data not shown). The extent of DJ_H recombination mediated by the wild-type and mutant proteins appeared roughly proportional to their activity on the cotransfected plasmid substrate when highly expressed with pEBB (Figs. 4B and 5A). At the level of RAG-1 expression supported by CDM8, however, it was apparent that we had reached the detection limit of the DJ_H PCR assay for the mutant RAG-1 alleles and thus could not establish a direct relationship between the extent of plasmid and endogenous recombination relative to R1WT in this range of RAG-1 expression.

The Severe Effect of N-terminal Mutations Is Evident When the Recombination Assay Is in a RAG-1 Dose-Responsive Range

When we measured plasmid recombination activity using the pEBB-RAG-1 vectors, we observed that the relative activities of the mutant RAG-1 alleles were higher than the values obtained using the equivalent amount of CDM8 vectors (Figures 4B and 5A). For example, NΔ1–384 was approximately one-third less active than R1WT when both were expressed from pEBB (Figs. 4B lanes 4 and 8; Figure 5A). Within this range of RAG-1 expression, however, transfection of 10-fold less pEBB-RAG-1 resulted in only a 3-fold decrease in recombination activity (Figure 4B, lanes 3 and 4, and Figure 5A). In contrast, the recombination assay was more directly proportional to the RAG-1 dose at a lower range of RAG-1 expression. We observed that transfection of 5-fold less CDM8-R1WT resulted in an approximately 3- to 4-fold decrease in recombination activity (Figure 5A). Furthermore, the amount of plasmid recombination supported by pEBB-R1WT was 9 times more than CDM8-R1WT (Table 1, row 1), consistent with the lower amounts of RAG-1 produced by CDM8. It was therefore evident that recombination activity in 1-2 cells using CDM8-RAG1 vectors was proportional to the RAG-1 dose over a broad range. These results suggested that the critical parameter that revealed the importance of the RAG-1 N-terminus might be the sensitivity of the assay to RAG-1 concentrations. Thus, we examined this issue further.

Previous assays to test the effect of RAG mutations on recombination activity have been performed in fibroblasts by cotransfecting CDM8-RAG-1 and RAG-2 expression vectors. In those systems, values for the activity of N-terminal truncations were relatively high (Sadofsky et al., 1993; Silver et al., 1993; Kirch et al., 1996). We also obtained substantially higher values for the relative activities of all the mutants when tested in 3T3 cells under similar conditions (Figure 5A and data not shown). However, under these conditions, the recombination assay was insensitive to the amount of input RAG-1 (Table 1, row 2), in striking contrast to our observations in the Abelson line.

We tested whether increasing the amount of RAG-2 in 3T3 cells could achieve assay conditions that would be more sensitive to the amount of input RAG-1. We used the vector pEBB-RAG2, which yielded greater than 20-fold higher amounts of RAG-2 protein than CDM8-RAG2 (Figure 3D). This increased the amount of recombination in 3T3 cells using CDM8-RAG-1 by over 15-fold (Table 1, row 4). Therefore, recombination activity was limiting for RAG-2 in 3T3 cells. Even with expression of higher amounts of RAG-2, while the assay became sensitive to the introduction of less CDM8-RAG-1 (Figure 5A), it remained insensitive to higher amounts of RAG-1 (Table 1, row 3). Under these conditions, although the relative activity of NΔ1–384 remained high, the activity of the C113A and C2 point mutants did decrease relative to R1WT (Figure 5A). Thus, by higher expression of RAG-2, we could bring the 3T3 cell assay partially towards the behavior of the Abelson cell assay. Figure 5B is a summary representation of the relationship we believe exists between the RAG-1 concentration and the activity of mutants.

RAG-1 Is Functionally Conserved between Mice and Zebrafish

Having an in vivo complementation system for RAG-1, we examined the activity of other alleles of RAG-1. To study the evolutionary conservation of RAG-1, we tested whether zebrafish RAG-1 could complement the recombination deficiency in the RAG-$1^{-/-}$ Abelson line. The overall homology between the zebrafish (Danio rerio) and murine RAG-1 proteins is among the lowest between RAG-1 genes (Willett et al., 1997); most of the differences are in the N-terminal region, which shares only 37% amino acid identity and has numerous spacing differences. Thus we first concentrated on this area—

Table 1. Comparison of the Sensitivity of Plasmid Recombination to the Introduction of High (pEBB) and Low (CDM8) Amounts of RAG-1 or RAG-2 under Different Transfection Conditions

		RAG		
Row	Cell Line	Constant	Variable (pEBB or CDM8)	Relative Recombination Activity with pEBB (CDM8 = 1.0)
1	1-2	Endogenous RAG-2	RAG-1	9.0 ± 4.1
2	3T3	CDM8 RAG-2	RAG-1	2.2 ± 0.87
3	3T3	pEBB RAG-2	RAG-1	1.63, 1.86
4	3T3	CDM8 RAG-1	RAG-2	15.6 ± 4.75

The amount of plasmid recombination supported by CDM8-RAG-1 or RAG-2 in the last column is defined as 1.0. For example, in row 1, in 1-2 cells that contain endogenous RAG-2, the relative recombination activity supported by pEBB-R1WT is on average 9-fold higher than CDM8-R1WT. In row 3, the values from two independent experiments are shown.

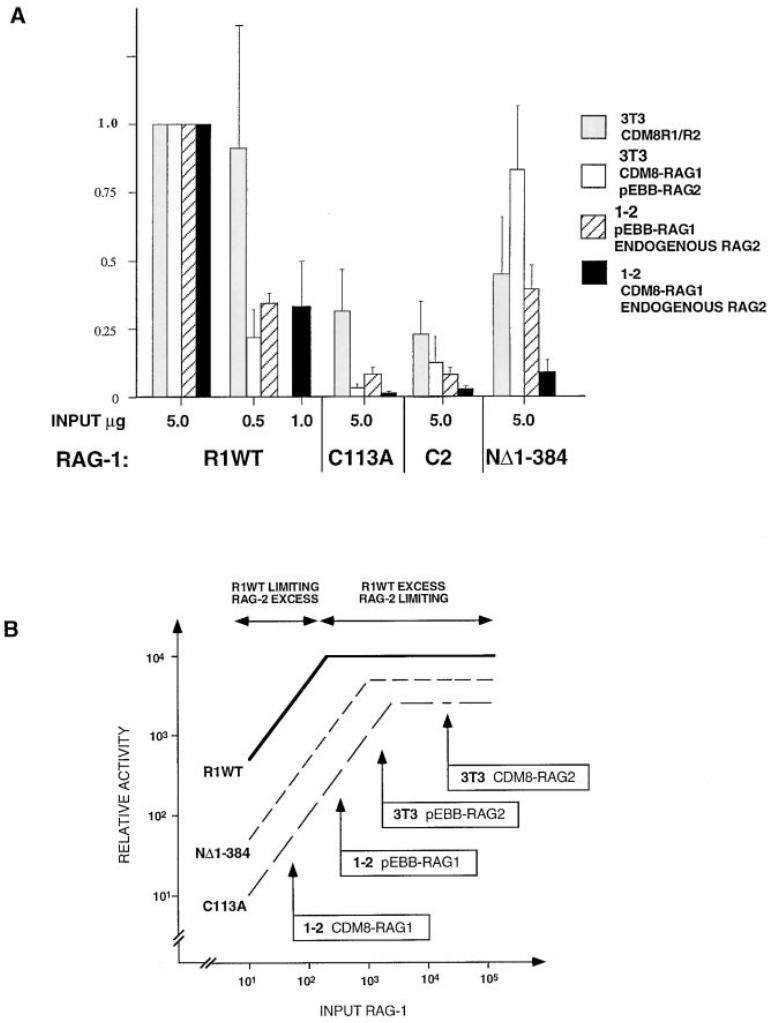

Figure 5. Comparison of the Relative Activities of a Panel of RAG-1 Alleles under Different Transfection Conditions

(A) Bar graph comparing the relative activities of RAG-1 alleles as measured under different transfection conditions. The activity for R1WT under each condition is defined as 1.00 and represents a relative, not absolute, amount of recombination. The identity of the RAG-1 allele and the amount transfected is indicated below each bar. The cell line and the expression vectors used are indicated in the bar graph legend.

(B) Graphic modeling of relative recombination activity to RAG-1 concentration. Based on observations of recombination in 1-2 cells, both axes are logarithmic and each tic represents an order of magnitude of arbitrary units. The arrows indicate the approximate position in the curve in which recombination activity is measured under the specified conditions. Since this graph represents relative, not absolute, recombination activity, and is a function of the amount of RAG-1, the absolute amounts of recombination under both 3T3 cell conditions cannot be directly compared to each other or to that in 1-2 cells. However, this graph can be used for direct comparison of recombination under both 1-2 conditions.

which could reveal the functional importance of the cysteine elements and the other conserved elements in the N-terminal region—by creating a chimeric RAG-1 molecule composed of the first 404 amino acids of zebrafish RAG-1 fused in-frame to the core domain of murine RAG-1. The activity of the chimeric protein, termed ZM-R1, was similar to that of the full-length murine protein and was greater than that of the NΔ1–384 allele in 1-2 cells (Figure 6A). Therefore, the N-terminal domain of zebrafish RAG-1 can functionally replace the murine domain, and the differences between the alleles of the two species highlights the importance of the elements that are conserved.

We then studied whether the functional conservation extended to the full-length zebrafish protein. Since the zebrafish RAG-1 coding region is interrupted by introns, we constructed an artificial zebrafish RAG-1 cDNA by multiple overlap PCR using a zebrafish genomic fragment as a template. Western blot analysis with a rabbit anti-mouse RAG-1 polyclonal serum raised to a conserved region within the core domain revealed the zebrafish protein with a band intensity about 3- to 5-fold less than R1WT (Figure 3D). Because we do not know the degree of cross-reactivity, this is a minimum estimate of the expression level. The zebrafish RAG-1 activated recombination of both plasmid and endogenous immunoglobulin recombination in the murine Abelson line, albeit at reduced levels (Figures 6A and 6B). Thus, a remarkable degree of RAG-1 function has been conserved between teleosts and mammals, representing 400 million years of divergence (Marchalonis, 1977). Although RAG-2 is only 53% conserved between these species, and the conservative elements are scattered evenly (Willett et al., 1997), the murine RAG-2 protein apparently can interact functionally with the zebrafish RAG-1 protein.

Discussion

We have established a physiological assay system using a RAG-1 deficient Abelson line to evaluate a panel of exogenous RAG-1 alleles for their ability to recombine plasmid and natural endogenous immunoglobulin substrates. Using this assay we have revealed a significant contribution of the catalytically dispensible N-terminal portion of RAG-1 to recombination activity. We have

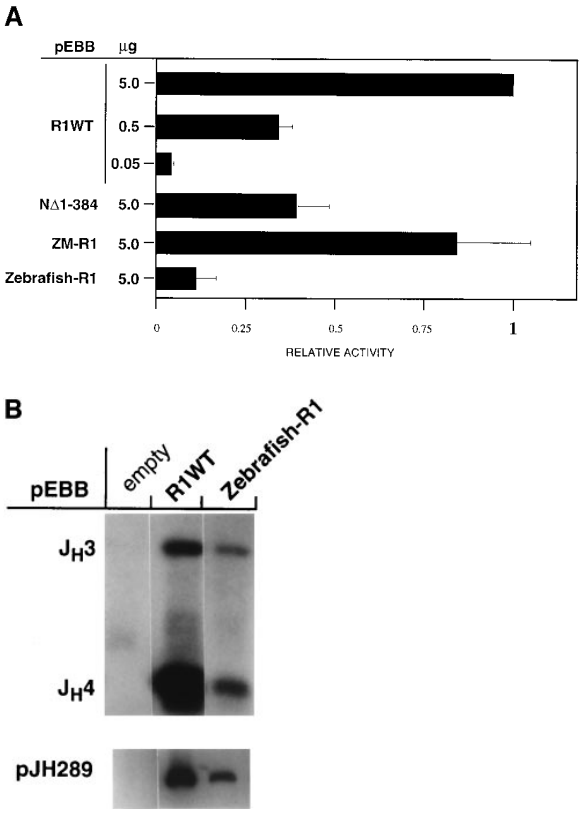

Figure 6. Chimeric and Full-Length Zebrafish RAG-1 Can Complement the V(D)J Recombination Deficiency in the RAG-1$^{-/-}$ Abelson Line 1-2

(A) Bar graph showing the relative activities of the indicated RAG-1 alleles expressed from pEBB using pJH289 as a readout for recombination activity. The amount of plasmid transfected is indicated to the left of each bar. "ZM-R1" is the N-terminal zebrafish/C-terminal murine RAG-1 chimera.

(B) Autoradiograph of DJ_H and pJH289 recombination PCR products from 1-2 cells transiently transfected with the indicated RAG-1 allele. In this particular PCR amplification, only bands representing J_H3 and J_H4 rearrangements were visualized, although all four products have collectively been observed from these samples.

defined a new potential zinc-binding motif, the D/EHC$_2$, and have shown that conserved cysteines within it are important functional components of the N-terminal portion of RAG-1. In the Abelson line, we found that the core domain of RAG-1 is sufficient to mediate rearrangement of natural immunoglobulin target sites and that zebrafish RAG-1 can complement the murine RAG-1 deficiency, providing the first description of the effect of mutations within RAG-1 on endogenous immunoglobulin recombination.

Our results show that the N-terminal region can contribute significantly to the V(D)J recombination activity of RAG-1. Under the appropriate conditions, the decrease in recombination activity relative to R1WT was over 10-fold for the N-terminal deletion mutant NΔ1–384 and more than 50-fold for the D/EHC$_2$ point mutants C1 and C2. To reveal fully the consequences of the N-terminal mutations, the recombination assay had to be within a RAG-1 dose-responsive range. This condition was fulfilled after the transient introduction of the CDM8-RAG-1 vectors into the the RAG-1$^{-/-}$ Abelson line and could be approached in fibroblasts by expressing high amounts of RAG-2. In the Abelson line, substantially more RAG-1 N-terminal mutant was required to achieve the same amount of recombination activity as R1WT (Figures 4B and 5A and as summarized in Figure 5B). Therefore, measurements of the recombination activity of the RAG-1 mutant proteins under these conditions provides an appropriate indication of their relative specific activities for recombination.

We began this study with the notion that there might be different requirements for segments of RAG-1 in lymphoid cells compared to fibroblasts or with recombination of endogenous DNA rather than the DNA of artificial constructs. We conclude, contrary to these expectations, that there are no evident differences in the assays we have performed. Rather, we find evidence for the same dependencies in all situations as long as recombination is limiting for RAG-1. Although we originally identified the phenotype of the N-terminal mutants in the Abelson line, the dependence on N-terminal elements we observed in this study was not cell line specific and thus apparently was not due to differences in the steady-state amounts or localization of RAG proteins or the presence of other tissue-specific factors or modifications of RAG-1.

The new and functionally important cysteine-containing elements, denoted D/EHC$_2$, within the N-terminal region of RAG-1 do not strictly contain canonical features used to classify other cysteine-rich motifs (Schwabe and Klug, 1994; Aasland et al., 1995). By virtue of containing the CxxC element, they resemble a portion of many zinc-binding motifs. The compact spacing of the putative coordinating amino acids in the D/EHC$_2$ also resembles the zinc-binding domain of retroviral nucleocapsid proteins (Myers et al., 1993). Although these domains in other proteins are involved in nucleic acid binding, we have no evidence that such an activity is associated with the D/EHC$_2$ elements. Instead, we suggest that the D/EHC$_2$ element bears more similarities to the zinc-binding H_2C_2 motif of retroviral integrase (IN), which is located within the N-terminal 100 amino acids of that protein (reviewed by Skalka, 1993). The IN H_2C_2 contains a CxxC element and, as we postulate for the cysteines in D/EHC$_2$, are the C-terminal zinc-coordinating amino acids (Bushman et al., 1993; Zheng et al., 1996). A comparison of H_2C_2 from many retroviruses indicated that there is some degree of flexibility in the the spacing between the histidine residues, and their distance from the CxxC is variable (Khan et al., 1991). Acidic residues have been postulated to facilitate Zn^{2+} coordination by IN, since only one histidine was shown to be essential for Zn^{2+} coordination (Burke et al., 1992); furthermore, acidic residues have been shown to directly coordinate Zn^{2+} in growth hormone (Cunningham et al., 1991). Such relationships strongly imply that these residues in the D/EHC$_2$s may also coordinate zinc and thus serve a structural role as they do in IN.

Other common properties of the N-terminal domains of RAG-1 and retroviral IN suggest how the N-terminus may contribute to RAG-1 activity. The H_2C_2 element plays a critical role in the activity of the IN protein (Khan et al., 1991; Vincent et al., 1993; Vink et al., 1993; Bushman and Wang, 1994); however, the H_2C_2 motif does not

directly participate in catalysis or DNA binding per se. Rather, its contribution to IN activity correlates with its ability to promote multimerization of IN, thereby allowing IN to efficiently cleave retroviral ends in *trans* (Engelman et al., 1993; van Gent et al., 1993; Donzella et al., 1996; Zheng et al., 1996). Point mutations of the histidine or cysteine residues within the H_2C_2 motif are more deleterious than truncations, perhaps because of the resultant inappropriate folding of the domain, which may interfere with an associative capability of the catalytic core (Donzella et al., 1996; Jonsson et al., 1996). These properties of the IN mutant proteins resemble what we have observed with RAG-1. First, we know that the N-terminal region is not required for DNA binding or catalysis per se because the core is functional in both respects. Thus, the core must have an independent capability to promote synapsis-dependent cleavage and recombination. Second, the activity of the RAG-1 C1 and C2 point mutants was significantly lower than NΔ1–384 under the appropriate conditions, suggesting that the point mutations were structurally disruptive and could thereby render the N-terminal region inhibitory. Third, the Zn-RING and ZnA together form homodimers in vitro, providing a precendent for a dimerization activity associated with this domain (Rodgers et al., 1996). Interestingly, an allele of RAG-1 that contains an internal deletion of both the homeodomain and the Zn-RING has been created (Kirch et al., 1996); it had previously been shown that N-terminal truncations that extend into the homeodomain render the remaining RAG-1 protein inactive (Sadofsky et al., 1993; Silver et al., 1993). However, this allele was activated to detectable levels when fused to the remainder of the N-terminal region that contained the D/EHC_2 elements (Kirch et al., 1996). In addition, the D/EHC_2 may explain differences in the nuclear localization properties of some BR mutations in full-length RAG-1 as compared to those within glutathione S-transferase –RAG-1-peptide fusion proteins that lacked the C1 element (Spanopoulou et al., 1995). These observations suggest that the C1 and C2 elements may define a structurally stable subdomain of the N-terminal region and that perhaps, like H_2C_2, they are involved in enhancing or stabilizing multimerization mediated by the core RAG-1 protein required for recombination. Although we suggest a zinc-binding and multimerization function for these elements, a formal demonstration awaits further biochemical analyses.

Our data show the strongest dependence on elements within the N-terminus when RAG-1 is limiting. Although the precise functional stoichiometries of active RAG proteins in vivo are not known, it is noteworthy that supplementing nuclear extracts from RAG-$1^{+/+}$, temperature-sensitive v-*abl*–transformed lines with recombinant RAG-1 protein stimulated in vitro cleavage activity present in the extract (van Gent et al., 1995). This is consistent with the contention that RAG-1 is limiting in this system despite the robust expression of both RAG-1 and RAG-2 in the nuclei induced by the inactivation of the v-*abl*^{ts} allele. It seems likely that the N-terminus of RAG-1 provides an interaction function that helps formation of RAG-1/RAG-2 complexes of appropriate stoichiometery, but in its absence a sufficient amount of active complexes can form when RAG-1 is present in excess. These higher-order RAG protein complexes may contain more RAG-2 than RAG-1, as has been suggested by immunoprecipation and sedimentation centrifugation studies (Leu and Schatz, 1995; Spanopoulou et al., 1995).

Previous assays have not revealed a strong contribution of portions of RAG-1 outside of the core domain to recombination in vivo. Specifically, the rearrangement of plasmid substrates or integrated retroviral substrates in fibroblasts has been only weakly dependent on the the N-terminal domain (Sadofsky et al., 1993; Silver et al., 1993; Kirch et al., 1996). Under the conditions of those assays, however, it is likely that RAG-1 was saturating, and the differences we report here could not have been evident. In addition, we have noted that at very low RAG-2 levels, there is a modest inhibition of recombination activity by RAG-1, with a 10-fold dilution of RAG-1 supporting a higher level of recombination than undiluted material; under those conditions mutants could even appear more active than R1WT (C. A. J. R., unpublished data).

Current in vitro assays for RAG function that measure cleavage of oligonucleotides (McBlane et al., 1995; Cuomo et al., 1996; Ramsden et al., 1996; Hiom and Gellert, 1997), plasmids (van Gent et al., 1995; Cortes et al., 1996; Eastman et al., 1996; van Gent et al., 1996b) or RSSs within chromatin from purified nuclei (Stanhope-Baker et al., 1996) do not require the N-terminal domain of RAG-1. In fact, the removal of the N-terminal region has been used to expedite solubilization and purification of recombinant RAG-1 proteins for these assays. Furthermore, functional stoichiomentries of RAG-1 and RAG-2 in vitro have not been rigorously determined. Certain epitope tags, such as His_6, have been shown to partially rescue H_2C_2-deleted IN proteins, and in some cases this correlated with the ability of the tag to facilitate multimerization (Bushman and Wang, 1994). Therefore, the addition of protein tags to core RAG-1 could dampen a possible effect of the absence of the N-terminal region. In addition, in vitro cleavage analyses of IN have shown that a deletion or point mutation of H_2C_2 is evident only in the presence of Mg^{2+} and that Mn^{2+} can mask the deleterious effect of those mutations, perhaps by increasing the associative capability of the core IN, as has been suggested for mutants of MuA (Baker and Luo, 1994). Oligonucleotide cleavage by the cores of RAG-1 and RAG-2, which mechanistically resembles DNA cleavage by the IN and MuA proteins (van Gent et al., 1996a), is generally done in the presence of Mn^{2+}. Although the conditions are nonphysiological, cleavage of isolated oligonucleotide RSSs may demand strong or particular protein–protein interactions that are enhanced by Mn^{2+}. Therefore, modified experimental conditions may reveal an independent contribution of the N-terminal domain in vitro.

Because we have found that NΔ1–384 can support endogenous immunoglobulin rearrangement, it is evident that—contrary to our expectation—the N-terminal domain does not provide any unique mechanistic function required for V(D)J recombination in vivo. RAG-1 lacking the N-terminus is able to mediate synapsis and joining between chromosomal target sites that are orders of magnitude further apart than those present in

artificial substrates. In addition, NΔ1–384 is sufficient to bind to RSSs within chromatin. The accessibility of DNA binding sites within chromatin is highly regulated, and thus its rearrangement provides a stringent and physiological readout for RAG-1 function. As a DNA binding complex, the RAG proteins should be subject to constraints similar to those regulating the access of transcription factors to DNA organized in chromatin. By comparison, DNA binding sites within transiently transfected plasmids are usually constituitively accessible to DNA-binding proteins (for example, Archer et al., 1992). Our results are in good agreement with the recent observation that recombinant core RAG-1 protein can promote cleavage of RSSs within isolated nuclei when added to a nuclear extract in vitro (Stanhope-Baker et al., 1996). Our results also extend that capability to the complete recombination reaction in vivo, since it was not determined whether cleavage in that system was 12/23-rule dependent or synapsis dependent because the assays were performed in the presence of Mn^{2+}. Furthermore, we have shown that the truncated RAG-1 protein can perform these functions in the absence of any full-length RAG-1. Therefore, if interactions with other proteins are required to bind to DNA sites within chromatin, that capability should be contained within the core domain. For example, it has been reported that human immunodeficiency virus (HIV)–IN interacts with hSNF5, a member of the SWI/SNF complex (Kalpana et al., 1994); hSNF5 also increased the activity of IN in vitro. Similarly, the MuB protein targets the integration site of MuA, with which it directly interacts (reviewed by Baker, 1995).

It has been postulated that the RAGs were originally derived from a transposable element (Thompson, 1995). Given that the D/EHC_2 elements are so far idiosyncratic to RAG-1, as is the H_2C_2 of IN, these elements may also have been present in the ancestral RAG-1 protein. Collectively the domains we described have been functionally conserved between mice and zebrafish, which represents an approximately 400 million–year divergence (Marchalonis, 1977). Although the activity of the full-length zebrafish protein was reduced significantly compared to murine RAG-1, it still was able to support plasmid and endogenous recombination with murine RAG-2. Since it is active, it considerably narrows the amino acids critical for RAG-1/RAG-2 association and catalysis.

The transient transfection system has allowed us to circumvent potential obstacles in assessing RAG-1 function. For example, dysregulated or overexpression of RAG-1 may be detrimental to prolonged cellular viability (Rathbun et al., 1993 and C. A. J. R., unpublished data). This system can be used to establish conditions under which the activity of the N-terminal region or other portions of RAG-1 or RAG-2 would be evident in vivo, to provide a paradigm for in vitro recombination systems and to understand more completely the phenotype associated with the restoration of any RAG-1–deficient system with mutant RAG-1 alleles.

Experimental Procedures

Cell Lines

Bone marrow–derived Abelson lines from 1-month-old RAG-$1^{-/-}$ mice (Spanopoulou et al., 1994) were established by conventional methods (Rosenberg et al., 1975). The v-*abl* retrovirus used to infect bone-marrow cultures was produced from a helper-free retroviral packaging system (Pear et al., 1993) using the construct pGD-v-*abl* (Scott et al., 1991). These cells were maintained in RPMI (Gibco–BRL) plus 10% fetal calf serum (JRH Biosciences), 50 μM β-ME (Sigma), and Pen/Strep (Gibco–BRL). 3T3 cells were maintained in Dulbecco's modified Eagle's medium (DMEM) plus 10% donor bovine serum, and 293 cells in DMEM plus 10% fetal calf serum prior to transfection.

Generation of RAG-1 Mutants

All RAG cDNAs were expressed either from CDM8 (Seed and Aruffo, 1987) or pEBB (Mizushima and Nagata, 1990; Mayer et al., 1995). Truncations (except NΔ1–384) and the BR mutant BRI-PM (K142I/K143I) are described in Silver et al. (Silver et al., 1993) and were rederived from their original constructs into the wild-type murine (M2) RAG-1 allele (His609; Roman and Baltimore, 1996). Site-directed mutations of RAG-1 were performed as described (Roman and Baltimore, 1996). The identity of the cysteine-element point mutations are as follows: C1, C110A/C113A; C2, C175A/C178A; C3, C210A/C213A; CH, C266A/H270A; ZnA, C355A; and HR2, R156Q/R159Q. The identity of other BR mutants (Spanopoulou et al., 1995) are as follows and were the gift of E. Spanopoulou (Mount Sinai): ΔBRII, Δ217–230; BRII-PMα, M222–225/233–236; BRII-PMβ, M233–236.

We generated an artificial zebrafish RAG-1 cDNA using a zebrafish genomic DNA fragement that contained the RAG-1 gene (Willett et al., 1997; a gift of C. Willett and L. Steiner, Massachusetts Institute of Technology) as a template for multistep overlap PCR to join the three RAG-1 exons. To create the chimera, an FseI site was first generated in murine RAG-1 by the introduction of a silent point mutation at G390 (GGA to GGC). A fragment of DNA that contained codons 1–409 of zebrafish RAG-1 was then amplified using a 3′ oligonucleotide that contained the FseI site. This fragment replaced the analogous murine sequence such that the chimera comprised zebrafish aa 1–405 and murine aa 387–1040. The NΔ1–384 allele was also made from the RAG-1 FseI allele by excising a BamHI/FseI fragment that contained the entire N-terminal region and replacing it with an oligonucleotide that contained a BamHI cohesive end and an initiating methionine followed by codons 384 to the FseI end. The integrity of all constructs was confirmed by sequence analysis.

Immunological Detection of RAG Proteins

To detect RAG proteins, lysates from 293 cells transiently transfected with RAG-1 or RAG-2 constructs were resolved by electrophoresis through SDS–polyacrylamide gel electrophoresis and probed with either a mouse anti-murine RAG-1 monoclonal antibody (Spanopoulou et al., 1995; Pharmingen), a rabbit anti-murine RAG-1 polyclonal antibody (Santa Cruz), or a mouse anti-murine RAG-2 monoclonal antibody (Pharmingen). Secondary antibodies (Amersham) were developed by enhanced chemiluminescence (Sigma).

Transfection Protocols

DNA was introduced into Abelson lines by DEAE-dextran–mediated transfection. Five micrograms of CDM8 or pEBB vectors (± RAG-1) were cotransfected with 5 μg of pJH289 or 288 (Lieber et al., 1988). DNAs were mixed in a total volume of 0.95 ml TD buffer (137 mM NaCl, 5 mM KCl, 0.4 mM Na_2HPO_4, 0.5 mM $MgCl_2$, 0.7 mM $CaCl_2$, and 25 mM Tris [pH 7.4]), and then 50 μl of 10 mg/ml DEAE-dextran (Sigma; 5×10^5 average molecular weight units) was added. Next, 10^7 log-phase cells were washed once in TD, and the pellet was resuspended in the 1 ml of DNA/DEAE-dextran mix and incubated for 10 min at room temperature. Cells were then resuspended in 10 ml of complete medium, spun out, resuspended in fresh medium, and plated for 48 hr before harvest.

DNA was introduced into 3T3 cells by precipitation with calcium phosphate 12–18 hr after plating at 5×10^5 cells per 60 mm dish. Cells were shocked with 15% glycerol for 2 min between 3–4 hr after the deposition of the precipitate. Typically, either 1 μg of CDM8-RAG-2 and 5 μg of CDM8, or 5 μg of pEBB-RAG-2, was coprecipitated with 5 μg of RAG-1 vector (± RAG-1 insert) and 5 μg pJH289.

Recombination Assays

Transfected cell pellets were washed once in phosphate-buffered saline and lysed in 400 μl of PCR lysis buffer (Schlissel et al., 1991); 1 μl of this lysate was used for amplification of plamid recombination products in a 25 μl PCR assay (Roman and Baltimore, 1996). To determine relative activities, a dilution curve for the PCR amplification was generated for each transfection: the sample representing R1WT was serially diluted into the negative control, and the dilution series was amplified in parallel with the control and experimental samples. Values for band intensities of the radiolabeled PCR products were determined by PhosphorImaging (Molecular Dynamics) with R1WT set as 1.00. Within the range of our assay, band intensities decreased proportionally to the dilution of the R1WT sample.

For genomic amplifications, the PCR lysate was phenol-extracted, concentrated by precipitation, and resuspended into 50–100λ TE_{-1} (10 mM Tris [pH 8.0], EDTA 1 mM). Approximately 1/20th of the recovered DNA was used as a template for PCR. DJ_H PCR fragments were amplified using the primers D_{HL} and J_H4 as described (Schlissel et al., 1991), except 40 cycles were used. PCR products were detected by Southern analysis using the oligonucleotide DR_{218} (Zhu et al., 1996).

Acknowledgments

We thank Ben Chen and Drs. Anthony Koleske and Joel Pomerantz for comments on the manuscript; Ben Chen and Drs. Konstantina Alexandropoulos, Joshy Jacob, Ilana Stancovsky, Kalle Saksela, and Martin Scott for advice and discussions. We are grateful to Drs. Catherine Willett and Lisa Steiner for the zebrafish RAG genomic fragment and RAG-1 sequence prior to publication; Dr. Eugenia Spanopoulou for RAG-1 BR mutants; and Pam Svec for excellent technical assistance. C. A. J. R. was supported by a postdoctoral award from the American Cancer Society. S. R. C. was supported by National Institutes of Health (NIH) predoctoral training grant 5-T32-GM07287. D. B. is an American Cancer Society Scholar. Supported by NIH grant CA51462.

Received March 12, 1997; revised May 14, 1997.

References

Aasland, R., Gibson, T.J., and Stewart, A.F. (1995). The PHD finger: implications for chromatin-mediated transcriptional regulation. Trends Biol. *20*, 56–59.

Alt, F., Rosenberg, N., Lewis, S., Thomas, E., and Baltimore, D. (1981). Organization and reorganization of immunoglobulin genes in A-MuLV–transformed cells: rearrangement of heavy but not light chain genes. Cell *27*, 381–390.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.

Archer, T.K., Lefebvre, P., Wolford, R., and Hager, G.L. (1992). Transcription factor loading on the MMTV promoter: a bimodal mechanism for promoter activation. Science *255*, 1573–1576.

Baker, T.A. (1995). Bacteriophage Mu: a transposing phage that integrates like retroviruses. Virology *6*, 53–63.

Baker, T.A., and Luo, L. (1994). Identification of residues in the Mu transposase essential for catalysis. Proc. Natl. Acad. Sci. USA *91,* 6654–6658.

Bernstein R.M., Schluter, S.F., and Bernstein, H., and Marchalonis, J.J. (1996). Primordial emergence of the recombination activating gene 1 (RAG1): Sequence of the complete shark indicates homology to microbial integrases. Proc. Natl. Acad. Sci. USA *93,* 9454–9459.

Burke, C.J., Sanyal, G., Bruner, M.W., Ryan, J.A., LaFemina, R.L., Robbins, H.L., Zeft, A.S., Middaugh, C.R., and Cordingley, M.G. (1992). Structural implications of spectroscopic characterization of a putative zinc-finger peptide from HIV-integrase. J. Biol. Chem. *267*, 9639–9644.

Bushman, F.D., Engelman, A., Palmer, I., Wingfield, P., and Craigie, R. (1993). Domains of the integrase protein of human immunodeficiency virus type 1 responsible for polynucleotidyl transfer and zinc binding. Proc. Natl. Acad. Sci. USA *90*, 3428–3432.

Bushman, F.D., and Wang, B. (1994). Rous sarcoma virus integrase protein: mapping functions for catalysis and substrate binding. J. Virol. *68*, 2215–2223.

Chang, Y., Paige, C.J., and Wu, G. (1992). Enumeration and characterization of DJH structures in mouse fetal liver. EMBO J. *11*, 1891–1899.

Chen, H., Kalaitsidaki, M., Warren, A.C., Avramopoulos, D., and Antonarakis, S.E. (1993). A novel zinc finger cDNA with a polymorphic pentanucleotide repeat (ATTTT)n maps on human chromosome 19p. Genomics *15*, 621–625.

Cortes, P., Ye, Z.S., and Baltimore, D. (1994). RAG-1 interacts with the repeated amino acid motif of the human homologue of the yeast protein SRP1. Proc. Natl. Acad. Sci. USA *91*, 7633–7637.

Cortes, P., Weis-Garcia, F., Misulovin, Z., Nussenzweig, A., Lai, J.-S., Li, G., Nussenzweig, M.C., and Baltimore, D. (1996). *In vitro* V(D)J recombination: signal joint formation. Proc. Natl. Acad. Sci. USA *93*, 14008–14013.

Cunningham, B.C., Mulkerrin, M.G., and Wells, J.A. (1991). Dimerization of human growth hormone by zinc. Science *253*, 545–548.

Cuomo, C.A., Mundy, C.L., and Oettinger, M.A. (1996). DNA sequence and structure requirements for cleavage of V(D)J recombination signal sequences. Mol. Cell. Biol. *16*, 5683–5690.

Difilippantonio, M.J., McMahan, C.J., Eastmann, Q.M., Spanopoulou, E., and Schatz, D.G. (1996). RAG1 mediated signal sequence recognition and recruitment of RAG2 in V(D)J recombination. Cell *87*, 253–262.

Donzella, G.A., Jonsson, C.B., and Roth, M.J. (1996). Coordinated disintegration reactions mediated by moloney murine leukemia virus integrase. J. Virol. *70*, 3909–3921.

Eastman, Q.M., Leu, T.M.J., and Schatz, D.G. (1996). Initiation of V(D)J recombination in vitro obeying the 12/23 rule. Nature *380*, 85–88.

Engelman, A., Bushman, F.D., and Craigie, R. (1993). Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complex. EMBO J. *12*, 3269–3275.

Fuschiotti, P., Harindranath, N., Mage, R.G., McCormack, W.T., Dhanarajan, P., and Roux, K.H. (1993). Recombination activating genes-1 and −2 of the rabbit: cloning and characterization of germline and expressed genes. Mol. Immunol. *30*, 1021–1032.

Gellert, M. (1996). A new view of V(D)J recombination. Genes Cells *1*, 269–275.

Gilks, C.B., Bear, S.E., Grimes, H.L., and Tsichlis, P.N. (1993). Progression of interleukin-2 (IL-2)-dependent T cell lymphoma lines to IL-2-independent growth following activation of a gene (Gfi-1) encoding a novel zinc finger protein. Mol. Cell. Biology *13*, 1759–1768.

Hendrickson, E.A., Qin, X.Q., Bump, E.A., Schatz, D.G., Oettinger, M., and Weaver, D.T. (1991). A link between double-strand break-related repair and V(D)J recombination: the scid mutation. Proc. Natl. Acad. Sci. USA *88*, 4061–4065.

Hiom, K., and Gellert, M. (1997). A stable RAG1-RAG2-DNA complex that is active in V(D)J cleavage. Cell *88*, 65–72.

Jackson, S.P., and Jeggo, P.A. (1995). DNA double-strand break repair and V(D)J recombination: involvement of DNA-PK. Trends Biol. *20*, 412–415.

Jonsson, C.B., Donzella, G.A., Gaucan, E., Smith, C.M., and Roth, M.J. (1996). Functional domains of Moloney Murine Leukemia Virus integrase defined by mutation and complementation analysis. J. Virol. *70*, 4585–4597.

Kalpana, G.V., Marmon, S., Wang, W., Crabtree, G., and Goff, S.P. (1994). Binding and stimulation of HIV-integrase by a human homolog of yeast trascription factor SNF5. Science *266*, 2002–2006.

Khan, E., Mack, J.P.G., Katz, R.A., Kulkosky, J., and Skalka, A.M. (1991). Retroviral integrase domains: DNA binding and the recognition of LTR sequences. Nucl. Acids Res. *19*, 851–860.

Kirch, S.A., Sudarasanam, P., and Oettinger, M.A. (1996). Regions of RAG1 proteins critical for V(D)J recombination. Eur. J. Immunol. *26*, 886–891.

Leu, T.M., and Schatz, D.G. (1995). RAG-1 and RAG-2 are components of a high-molecular-weight complex, and association of

RAG-2 with this complex is RAG-1 dependent. Mol. Cell. Biol. *15,* 5657–5670.

Lieber, M.R., Hesse, J.E., Lewis, S., Bosma, G.C., Rosenberg, N., Mizuuchi, K., Bosma, M.J., and Gellert, M. (1988). The defect in murine severe combined immune deficiency: joining of signal sequences but not coding segments in V(D)J recombination. Cell *55*, 7–16.

Marchalonis, J.J. (1977). Immunity in Evolution (Cambridge, MA: Harvard University Press).

Mayer, B.J., Hirai, H., and Sakai, R. (1995). Evidence that SH2 domains promote processive phosphorylation by protein-tyrosine kinases. Curr. Biol. *5*, 296–305.

McBlane, J.F., van, G.D., Ramsden, D.A., Romeo, C., Cuomo, C.A., Gellert, M., and Oettinger, M.A. (1995). Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. Cell *83*, 387–395.

Mizushima, S., and Nagata, S. (1990). pEF-BOS, a powerful mammalian expression vector. Nucl. Acids Res. *18*, 5322.

Myers, G., Wain-Hobson, S., Korber, B., Smith, R.F., and Pavlakis, G.N., eds. (1993). Human Retroviruses as AIDS 1993 I-II (Los Alamos, NM: Theoretical Biology and Biophysics).

Oettinger, M.A., Schatz, D.G., Gorka, C., and Baltimore, D. (1990). RAG-1 and RAG-2, adjacent genes that synergystically activate V(D)J recombination. Science *248*, 1517–1523.

Oettinger, M.A. (1996). Cutting apart V(D)J recombination. Curr. Opin. Genet. *6*, 141–145.

Passananti, C., Felsani, A., Caruso, M., and Amati, P. (1989). Mouse genes encoding 'zinc-finger' containing proteins: characterization and expression in differentiating cells. Proc. Natl. Acad. Sci. USA *86*, 9417–9421.

Pear, W.S., Nolan, G.P., Scott, M.L., and Baltimore, D. (1993). Production of high-titer helper-free retroviruses by transient transfection. Proc. Natl. Acad. Sci. USA *90*, 8392–8396.

Ramsden, D.A., McBlane, J.F., van Gent, D.C., and Gellert, M. (1996). Distinct DNA sequence and structure requirements for the two steps of V(D)J recombination signal cleavage. EMBO J. *15,* 3197–3206.

Rathbun, G., Oltz, E.M., and Alt, F.W. (1993). Comparison of RAG gene expression in normal and transformed precursor lymphocytes. Int. Immunol *5*, 997–1000.

Rodgers, K.K., Bu, Z., Fleming, K., Schatz, D.G., Engelman, D.M. and Coleman, J.E. (1996). A zinc-binding comain involved in the dimerization of RAG1. J. Mol. Biol. *260*, 70–84.

Roman, C.A.J., and Baltimore, D. (1996). Genetic evidence that the RAG1 protein directly participates in V(D)J recombination through substrate recognition. Proc. Natl. Acad. Sci. USA *93*, 2333–2338.

Rosenberg, N., Baltimore, D., and Scher, C.D. (1975). In vitro transformation of lymphoid cells by Abelson murine leukemia virus. Proc. Natl. Acad. Sci. USA *72*, 1932–1936.

Sadofsky, M.J., Hesse, J.E., McBlane, J.F., and Gellert, M. (1993). Expression and V(D)J recombination activity of mutated RAG-1 proteins. Nucl. Acids Res. *21*, 5644–5650.

Saurin, A.J., Borden, K.L.B., Brody, M.N., and Freemont, P.S. (1996). Does this have a familiar RING? Trends Biol. *21*, 208–214.

Schlissel, M.S., Corcoran, L.M., and Baltimore, D. (1991). Virus-transformed pre-B cells show ordered activation but not inactivation of immunoglobulin gene rearrangement and transcription. J. Exp. Med. *173*, 711–720.

Schwabe, J.W.R., and Klug, A. (1994). Zinc mining for protein domains. Struct. Biol. *1*, 345–349.

Scott, M.L., Etten, R.A.V., Daley, G.Q., and Baltimore, D. (1991). v-abl causes hematopoietic disease distinct from that caused by bcr-abl. Proc. Natl. Acad. Sci. *88*, 6506–6510.

Seed, B., and Aruffo, A. (1987). Molecular cloning of the CD2 antigen, the T-cell erythrocyte receptor, by a rapid immunoselection procedure. Proc. Natl. Acad. Sci. USA *84*, 3365–3369.

Silver, D.P., Spanopoulou, E., Mulligan, R.C., and Baltimore, D. (1993). Dispensible sequence motifs in the RAG-1 and RAG-2 genes for plasmid V(D)J recombination. Proc. Natl. Acad. Sci. USA *90*, 6100–6104.

Skalka, A.M. (1993). Retroviral DNA integration: lessons from transposon shuffling. Gene *135*, 175–182.

Spanopoulou, E., Roman, C.A.J., Corcoran, L.M., Schlissel, M.S., Silver, D., Nussenzweig, M., Shinton, S., Hardy, R., and Baltimore, D. (1994). Functional immunoglobulin transgenes guide ordered B-cell differentiation in RAG-1-deficient mice. Genes Dev. *8*, 1030–1042.

Spanopoulou, E., Cortes, P., Shih, C., Huang, C.-H., Silver, D.P., Svec, P., and Baltimore, D. (1995). Localization, interaction, and RNA binding properties of the V(D)J recombination-activating proteins RAG1 and RAG2. Immunity *3*, 715–726.

Spanopoulou, E., Zaitseva, F., Wang, F.-H., Santagata, S., Baltimore, D., and Panayotou, G. (1996). The homeodomain region of RAG-1 reveals the parallel mechanisms of bacterial and V(D)J recombination. Cell *87*, 263–276.

Stanhope-Baker, P., Hudson, K.M., Shaffer, A.L., Contantinescu, A. and Schlissel, M.S. (1996). Cell type-specific chromatin structure determines the targetting of V(D)J recombinase activity in vitro. Cell *85*, 887–897.

Thompson, C.B. (1995). New insights into V(D)J recombination and its role in the evolution of the immune system. Immunity *3*, 531–539.

van Gent, D.C., Vink, C., Groeneger, A.M.O., and Plasterk, R.H.A. (1993). Complemetation between HIV integrase proteins mutated in different domains. EMBO J. *12*, 3261–3267.

van Gent, D.C., McBlane, J.F., Ramsden, D.A., Sadofsky, M.J., Hesse, J.E., and Gellert, M. (1995). Initiation of V(D)J recombination in a cell-free system. Cell *81*, 925–934.

van Gent, D.C., Mizuuchi, K., and Gellert, M. (1996a). Similarities between initiation of V(D)J recombination and retroviral integration. Science *271*, 1592–1594.

van Gent, D.C., Ramsden, D.A., and Gellert, M. (1996b). The RAG1 and RAG2 proteins establish the 12/23 rule in V(D)J recombination. Cell *85*, 107–113.

Vincent, K.A., Ellison, V., Chow, S.A., and Brown, P.O. (1993). Characterization of human immunodeficiency virus type 1 integrase expressed in *Escherichia coli* and analysis of variants with amino-terminal mutations. J. Virol. *67*, 425–437.

Vink, C., Öude Groeneger, A.A.M., and Plasterk, R.H.A. (1993). Identification of the catalytic and DNA binding region of the human immunodeficiency virus type 1 integrase protein. Nucl. Acids Res. *21*, 1419–1425.

Weaver, D.T. (1995). What to do at an end: DNA double-strand-break repair. Trends Genet. *11*, 388–392.

Willett, C.E., Cherry, J.J., and Steiner, L.A. (1997). Characterization and expression of the recombination activating genes (*RAG-1* and *RAG-2*) of zebrafish. Immunogenetics *45*, 394–404.

Witzgall, R., O'Leary, E., Gessner, R., Ouellette, A.J., and Bonventre, J. V. (1993). Kid-1, a putative renal transcription factor: regulation during ontogeny and in response to ischemia and toxic injury. Mol. Cell. Biol. *13*, 1933–1942.

Zheng, R., Jenkins, T.M., and Craigie, R. (1996). Zinc folds the N-terminal domain of HIV-integrase, promotes multimerization, and enhances catalytic activity. Proc. Natl. Acad. Sci. *93*, 13659–13664.

Zhu, C., Bogue, M.A., Lim, D.S., Hasty, P., and Roth, D.B. (1996). Ku86-deficient mice exhibit severe combined immunodeficiency and defective processing of V(D)J recombination intermediates. Cell *86*, 379–389.